# STATISTICS AND EVIDENCE SYNTHESIS

## 1. What is evidence synthesis?

## 2. Why do we need it?

## 3. How to do it?

Sharon-Lise T. Normand
Harvard Medical School & Harvard School of Public Health

Silver Spring, MD: March 7, 2012

# EXAMPLES

| DEVICE | Esthetic Devices<br>Silicone Gel-Filled Breast Implants | Pediatric Therapeutic Devices<br>Vagel Nerve Stimulators | Adult Therapeutic Devices<br>Lap Band |
|---|---|---|---|
| **Indication** | Augmentation (≥ 22 yrs old) or reconstruction | Refractory epilepsy patients (approved in US for ≥ 12 years old) | Weight reduction in severely obese subjects (BMI ≥ 40 or BMI ≥ 35+comorbids) |
| **Device** | Silicone shell filled with silicone gel | Electrical pulses applied to left vagel nerve to send signals to brain; device = generator + leads | Silicone gastric band, access port, and tubing to connect port to band |
| **Alternatives** | Saline implant (≥ 18 yrs old) | Resective surgery; ketogenic diet | Behaviorial therapy; $R_x$; jaw-wiring, hypnosis; surgery (gastric bypass or vertical banded) |
| **Outcomes** | Bra cup size change, chest size, quality of life; safety = rash, wrinkling, malposition, leakage | No new AED drug **and** no "significant" dose change within 1-year; **and** ≥ 50% decrease from baseline in seizure frequency at 1-month | Weight loss; safety = digestive problems, band slippage or dilation; port displacement; port or tube leakage |
| **Challenges in Post Market Setting** | Patients **don't want** to be identified; case ascertainment problems | Premarket data are **sparse** and sometimes do not exist | **Indications expanded** to less obese patients |

# EVIDENCE SYNTHESIS

*Evidence synthesis involves the development of techniques to combine multiple sources of quantitative evidence.*

- this is beyond meta-analysis
- individual data, cohort data, etc.
- need to define **evidence**

# WHAT ARE THE SPECIFIC QUESTIONS?

- What do we **expect** to see in the postmarket setting for a new device?
  - use the PMA data alone?
  - identify target population (patients, providers, settings)
  - quantify the relative prevalence of covariates in PMA population and target population (to infer expected effectiveness)
  - outcomes completely missing for target population so can only calculate predictions
- What information should be used?
  - Several different outcomes (index by $m$)
  - Several different device manufacturers (index by $k$)
  - Several different patient groups (index by $j$)
  - Several data sources (index by $i$)

# BORROWING STRENGTH

Suppose $j = 1, 2, \cdots, n_k$ observations within device **group** $k$ for $k = 1, 2, \cdots, K$. Let $y_k$ = mean outcome for a patient group implanted with device $k$.

$$y_k \sim N(\alpha_k, \sigma_y^2) \text{ and } \alpha_k \sim N(\mu_\alpha, \sigma_\alpha^2) \tag{1}$$

For each device $k$, the conditional (posterior) estimate of $\alpha_k$ is

$$\hat{\alpha}_k \mid y, \mu_\alpha = \omega_k \mu_\alpha + (1 - \omega_k)\bar{y}_k \tag{2}$$

weighted average of the within-device outcomes ($\bar{y}_k$) and between-device model for outcomes ($\mu_\alpha$)

# EXAMPLE

## Total Hip Implants

| Device Characteristic | Device | Outcome | Evidence Sources | | | |
|---|---|---|---|---|---|---|
| | | | PMA | Literature | Registry | Claims* |
| Ceramic on Ceramic | 1 | HHS | √ | | | No UDI No HHS |
| | | Revision | √ | | | |
| | | AE | √ | | | |
| | 2 | HHS | √ | | | |
| | | Revision | √ | | | |
| | | AE | √ | | | |
| Metal on Metal | 3 | HHS | Not available | | | No UDI No HHS |
| | | Revision | | | | |
| | | AE | | | | |
| | 4 | etc | √ | | | |

# BORROWING STRENGTH

$k$ may be a class of device defined by characteristic of device or may denote the manufacturer

$$\left(\frac{\text{No}}{\text{Pooling}}\right) \; 0 \leq \; \omega_k \quad = \quad \left(\frac{\frac{\sigma_y^2}{n_k}}{\sigma_\alpha^2 + \frac{\sigma_y^2}{n_k}}\right) \quad \leq \; 1 \; \left(\frac{\text{Complete}}{\text{Pooling}}\right)$$

- **Sampling Variability:** $\frac{\sigma_y}{\sqrt{n_k}}$
- **Signal** $= \alpha_k$
- **Strength**: $\text{Var}(\alpha_k) = (1 - \omega_k)\frac{\sigma_y^2}{n_k} \; \leq \quad \text{Var}(\bar{y}_k)$

# ASSUMPTIONS

1. Exchangeability: conditional on $X$, ordering of (size) device effects is indistinguishable

$$\alpha_k \sim N(\mu_\alpha, \sigma^2_\alpha) \tag{3}$$

   - While the effects themselves may differ, we can consider them drawn from a **common** distribution
2. Coherence: assumptions to accommodate **indirect** comparisons
   - want effect of device $k$ on outcome to be the same, **regardless** of what it is compared against
   - $var(\xi_{kl})$ close to 0 where $\xi_{kl}$ is **change** in $\alpha_k$ when compared to $\alpha_l$
3. Variance components: prior distributions to characterize uncertainty in strengths of relationships

# CONCLUDING REMARKS: HOW RELATED TO 522 STUDIES

- Increasing need for prediction and extrapolation of both safety and effectiveness outcomes
  - To determine **expected** outcomes
  - To bolster inferences when pre-market data are sparse
- Increasing need for incorporation of **uncertainty**
  - Accommodated through probability distributions
  - Critically important in the analysis of surveillance studies
- Quantifying **evidence**
  - Need to measure the **evidence** for a hypothesis
  - Bayes factors provide quantitative measures of evidence